

WG#12: Translingual Information Management

Report Summary

- **Problem/Goal/ Technical Scope**
 - **Enable a Human User, Operating in Own Native Language(s), To Effectively Access, Utilize, and Correlate Multilingual, Multimodal Sources of Information Relevant to a Specified Problem Domain**
 - **Subtopics: Learning/Acquisition; Translingual/Multimodal Generation**
- **Relevant Disciplines / Technologies**
 - **Content-Based Information Retrieval, Extraction, & Correlation**
 - **Machine Translation**
 - **Natural Language Understanding & Meaning Representation**
 - **Recognition & Generation**

Working Panel: Cliff Weinstein (chair); Bruce Schatz; Ralph Weischedel, Robert Frederking

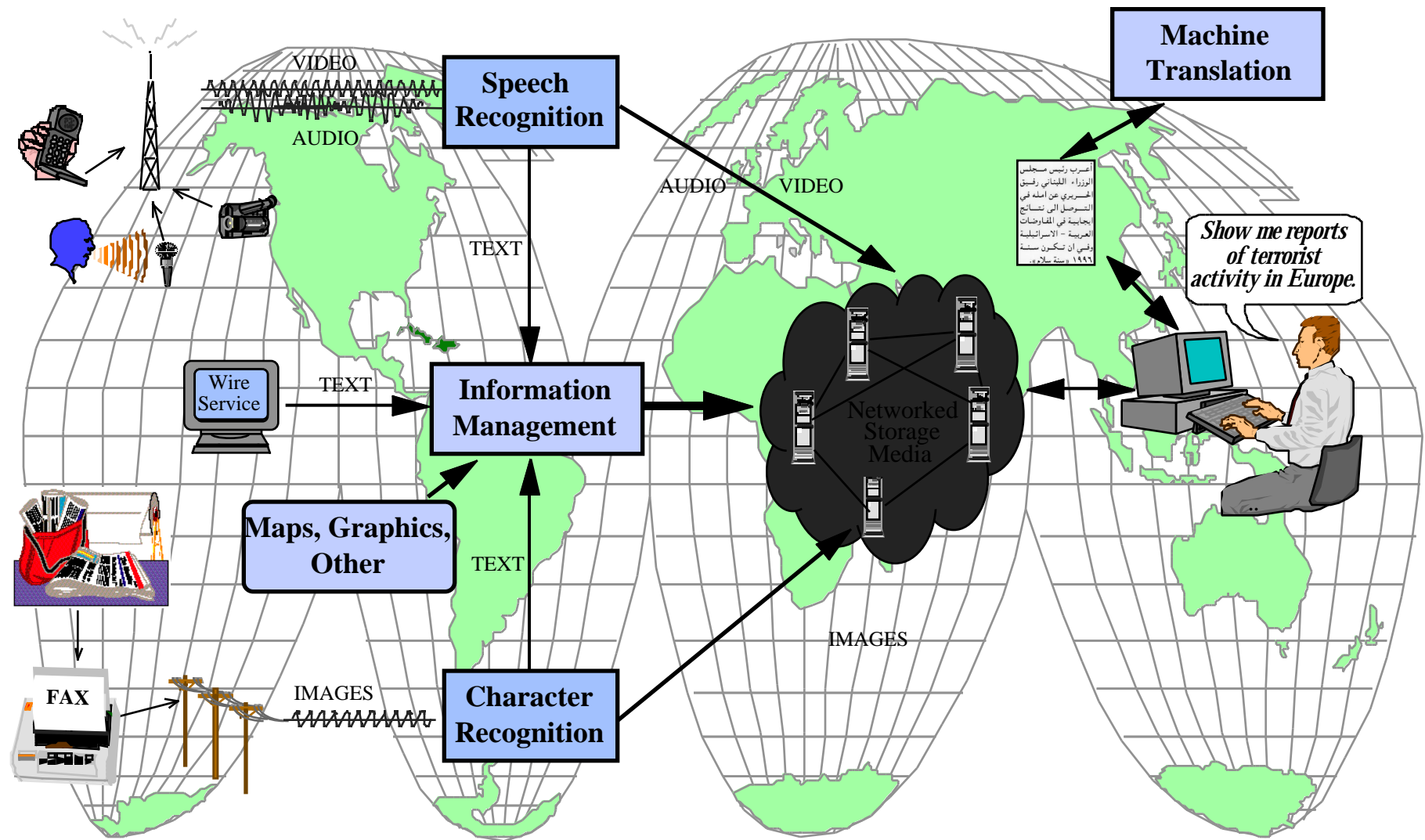
Sponsor: Ron Larsen; Co-Sponsor: Allen Sears

WG#12: Translingual Information Management

Sample Application Scenarios

- **Research Report for Intelligence**
 - **Mix of Open and Restricted Sources**
 - **Text, Images, Maps, Video, ...**
 - **Turnaround Time Range: 1-15 Days**
- **Battlefield Situation Analysis/Summary**
 - **Multilingual Sources, Some Poor Quality**
 - **Linguistic Info from Allies & Foes**
 - **Multimedia Including Sensor Data**
 - **Need to Fuse/Summarize Linguistic/Non-Linguistic Info**
 - **Turnaround Time Range: 5 sec - 30 minutes**

Translingual, Multi-modal Information Management



WG#12: Translingual Information Management

Related Programs & Ideas

- **Information Management**
- **Intelligent Collaboration**
- **Human Language Systems & Human/Computer Interaction**
- **TIPSTER**
- **Digital Library**
- **Translingual vs Multilingual**
- **Semantic Interoperability & Generalized Interlingua**
 - **Across Languages, Media, Representations (Graphics, Plans, ...)**

WG#12: Translingual Information Management Addressing the Challenges

- **Challenges**
 - **No Full Translingual Info Mgmt System Has Been Integrated or Tested**
 - **Make the Whole More effective than the sum of the Component Technologies**
 - **Efficient Generalization/Porting to New Languages & Domains**
 - **Need to Fuse Linguistic & Non-Linguistic Info**
 - **No Suitable Translingual Corpus or Tools**

WG#12: Translingual Information Management Technical & Strategic Approaches

- **NovelTechnical**
 - **Translingual Relevance Feedback**
 - **Knowledge-Based Interlingua**
 - **Translingual Concept Switching**
 - **Multilingual Wordnet**
- **Strategic**
 - **Can/Should Rapidly Build an Architecture and Prototype/Testbed**
 - **Establish a Baseline, Pinpoint the R&D Issues**
 - **Build Tools and Corpora**
 - **Focus on End-to-End, App-Oriented SysEval**

1996 DARPA ITO General PI Meeting, Dallas, TX

WG#12: Translingual Information Management Investment Strategy

- DARPA, Industry Support
 - Why DARPA? Too Long-Term for Industry; Special DoD Needs; Tool Sharing Needed
 - What other collaborations? International, e.g. CSTAR
- What if we did not do this?
 - It Won't Happen
 - *U.S. at Strategic Disadvantage* Because “They” Can & Will Get Info From Us
- Optimal Scale of Efforts
 - small vs large? Collaboration needed; mix? some small technology efforts

1996 DARPA ITO General PI Meeting, Dallas, TX

WG#12: Translingual Information Management

Other Issues Addressed

- **What is Novel vs Current Programs or Approaches**
- **Relevant Work by Industry or Others**
 - **Lots of Component Tech R&D**
 - **Translingual:TIPSTER, small research**
- **Relative Risk & Likelihood of Success of Proposed Ideas or Approaches**
 - **Some Success is likely, because we have almost nothing now**
 - **The full-blown problem is very challenging**